Intelligent Systems (AI-2)

Computer Science cpsc422, Lecture 26

March 19, 2021

422 big picture

StarAI (statistical relational AI)

Hybrid: Det +Sto

- *Prob CFG* — today
- *Prob Relational Models*
- *Markov Logics*

	Deterministic	Stochastic
Query	*Logics* *First Order Logics* *Ontologies* • Full Resolution • SAT	*Belief Nets* Approx. : Gibbs *Markov Chains and HMMs* Forward, Viterbi…. Approx. : Particle Filtering *Undirected Graphical Models* *Markov Networks* *Conditional Random Fields*
Planning		*Markov Decision Processes and Partially Observable MDP* • Value Iteration • Approx. Inference *Reinforcement Learning*

Applications of AI

Representation

Reasoning Technique

Linguistic Knowledge-Formalisms NLP Map

- Morphology
- Syntax
- Semantics
- Pragmatics Discourse and Dialogue

State Machines *(and prob. versions)*
(Finite State Automata, Finite State Transducers, *Markov Models, CRFs*)
Neural Models, Neural Sequence Modeling

Rule systems *(and prob. versions)*
(e.g., *(Prob.)* Context-Free Grammars)

Logical formalisms (First-Order Logics, *Prob. Logics*)

AI planners *(MDP Markov Decision Processes, Reinforcement learning)*

Machine Learning

Lecture Overview

- Recap English Syntax and Parsing
- Key Problem with parsing: Ambiguity
- Probabilistic Context Free Grammars (PCFG)
- Treebanks and Grammar Learning

Key Constituents: Examples

Head

NP -> N
NP -> Det N

(Specifier) X (Complement)

• Noun phrases (NP)	• (Det)	N	(PP)
	the	cat	on the tabl
• Verb phrases (VP)	• (Qual)	V	(NP)
	never	eat	a cat
• Prepositional phrases (PP)	• (Deg)	P	(NP)
	almost	in	the net
• Adjective phrases(AP)	• (Deg)	A	(PP)
	very	happy	about it
• Sentences (S)	• (NP)	(-)	(VP)
	a mouse	--	ate it

Context Free Grammar (CFG)

- **4-tuple** (non-term., term., productions, start)
- (N, Σ, P, S)
- P is a set of rules $A \rightarrow \alpha$; $A \in \mathbf{N}$, $\alpha \in (\Sigma \cup N)^*$

$N = \{X\ Y\}$ $\Sigma = \{a\ b\ c\}$ $P =$

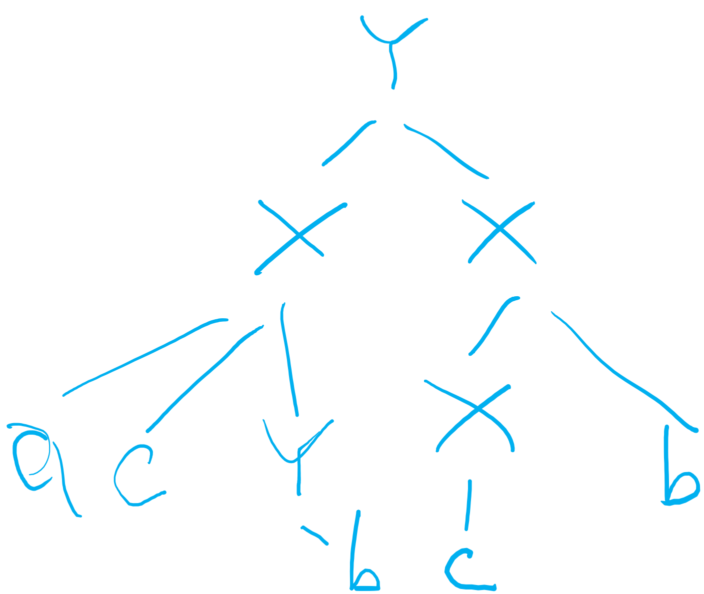
$X \rightarrow Xb$

$Y \rightarrow XX$

$X \rightarrow acY$

$X \rightarrow c$

$Y \rightarrow b$



Toy CFG Example for English

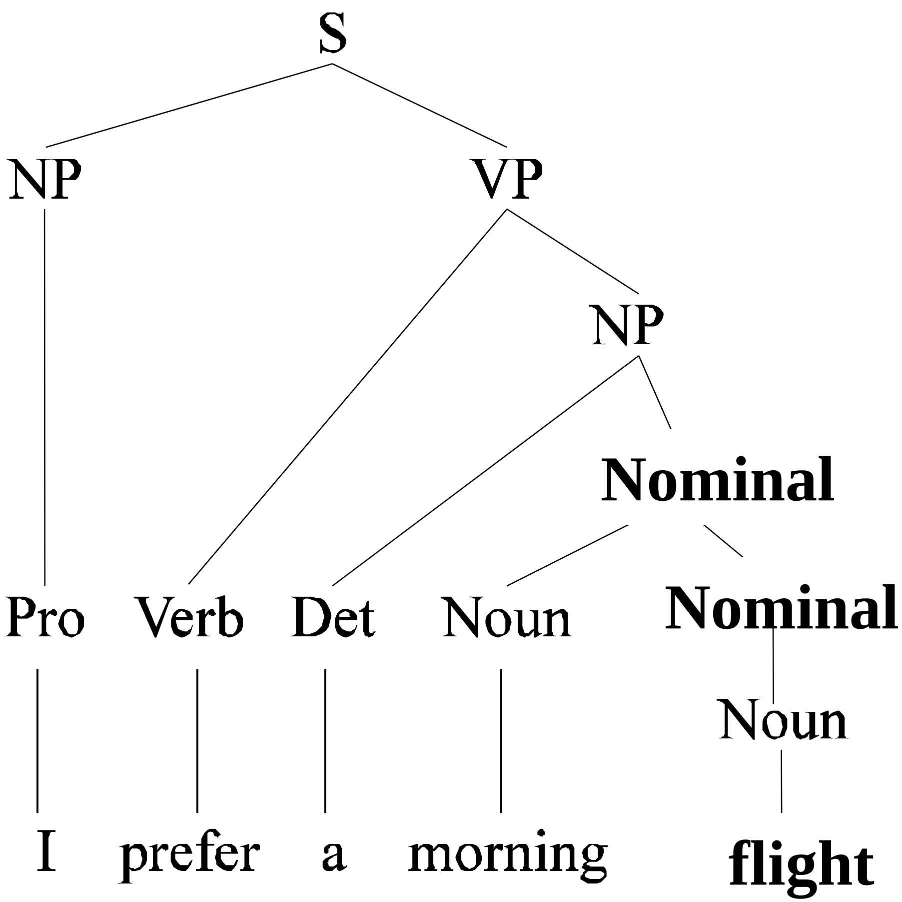
Grammar with example phrases

S	→	*NP VP*	I + want a morning flight
NP	→	*Pronoun*	I
	|	*Proper-Noun*	Los Angeles
	|	*Det Nominal*	a + flight
Nominal	→	*Noun Nominal*	morning + flight
	|	*Noun*	flights
VP	→	*Verb*	do
	|	*Verb NP*	want + a flight
	|	*Verb NP PP*	leave + Boston + in the morning
	|	*Verb PP*	leaving + on Thursday
PP	→	*Preposition NP*	from + Los Angeles

Lexicon

Noun	→	*flights* | *breeze* | *trip* | *morning* | ...
Verb	→	*is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective	→	*cheapest* | *non − stop* | *first* | *latest* | *other* | *direct* | ...
Pronoun	→	*me* | *I* | *you* | *it* | ...
Proper-Noun	→	*Alaska* | *Baltimore* | *Los Angeles* | *Chicago* | *United* | *American* | ...
Determiner	→	*the* | *a* | *an* | *this* | *these* | *that* | ...
Preposition	→	*from* | *to* | *on* | *near* | ...
Conjunction	→	*and* | *or* | *but* | ...

Derivations as Trees



$S \rightarrow NP\ VP$

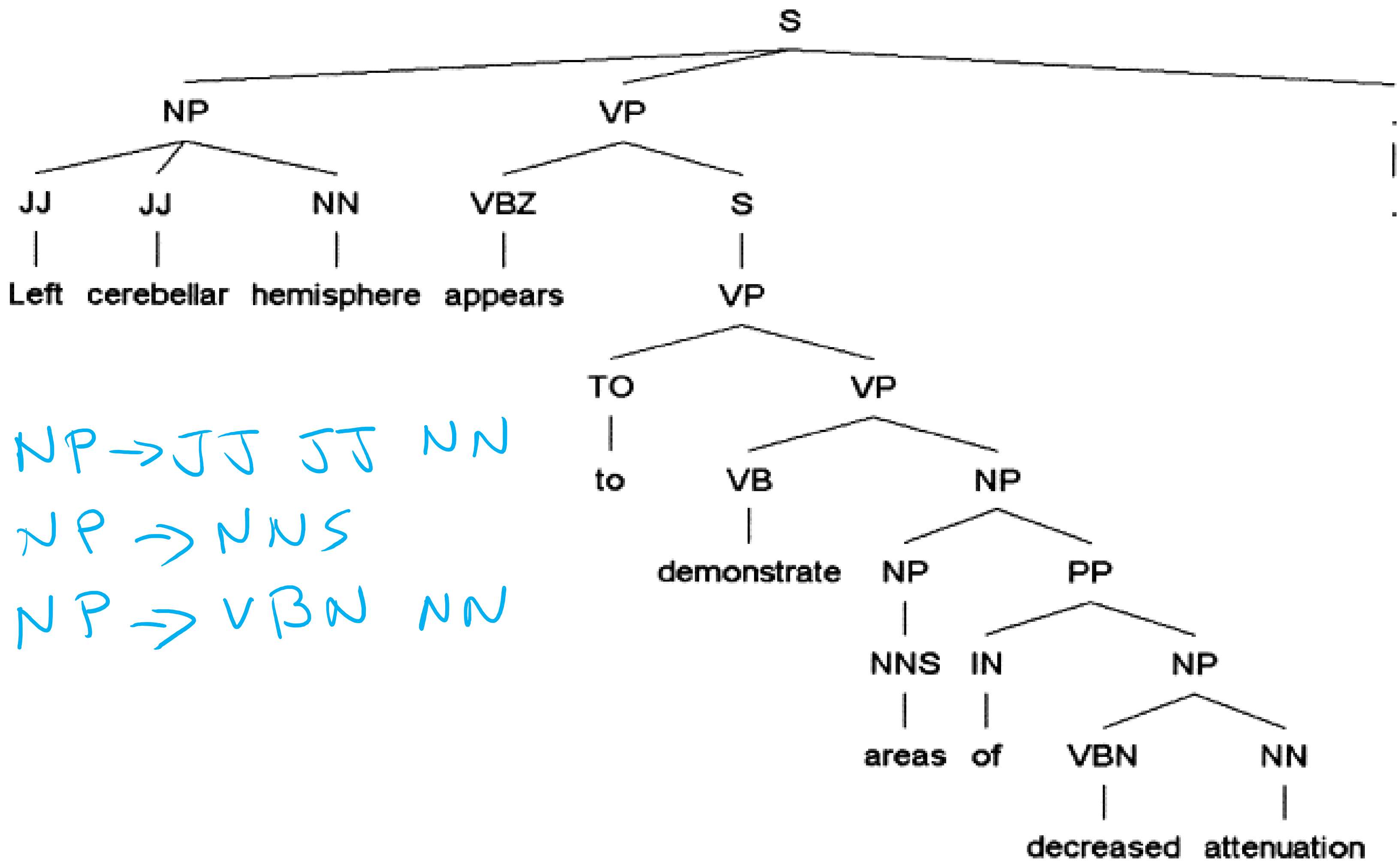
$NP \rightarrow Pronoun$
$\quad\ |\ Proper\text{-}Noun$
$\quad\ |\ Det\ Nominal$

$\text{Nominal} \rightarrow Noun\ Nominal$
$\quad\ |\ Noun$

$VP \rightarrow Verb$
$\quad\ |\ Verb\ NP$
$\quad\ |\ Verb\ NP\ PP$
$\quad\ |\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

Example of relatively complex parse tree



NP → JJ JJ NN
NP → NNS
NP → VBN NN

Journal of the American Medical Informatics Association, 2005, Improved Identification of Noun Phrases in Clinical Radiology Reports Using a High-Performance Statistical Natural Language Parser Augmented with the UMLS Specialist Lexicon

Lecture Overview

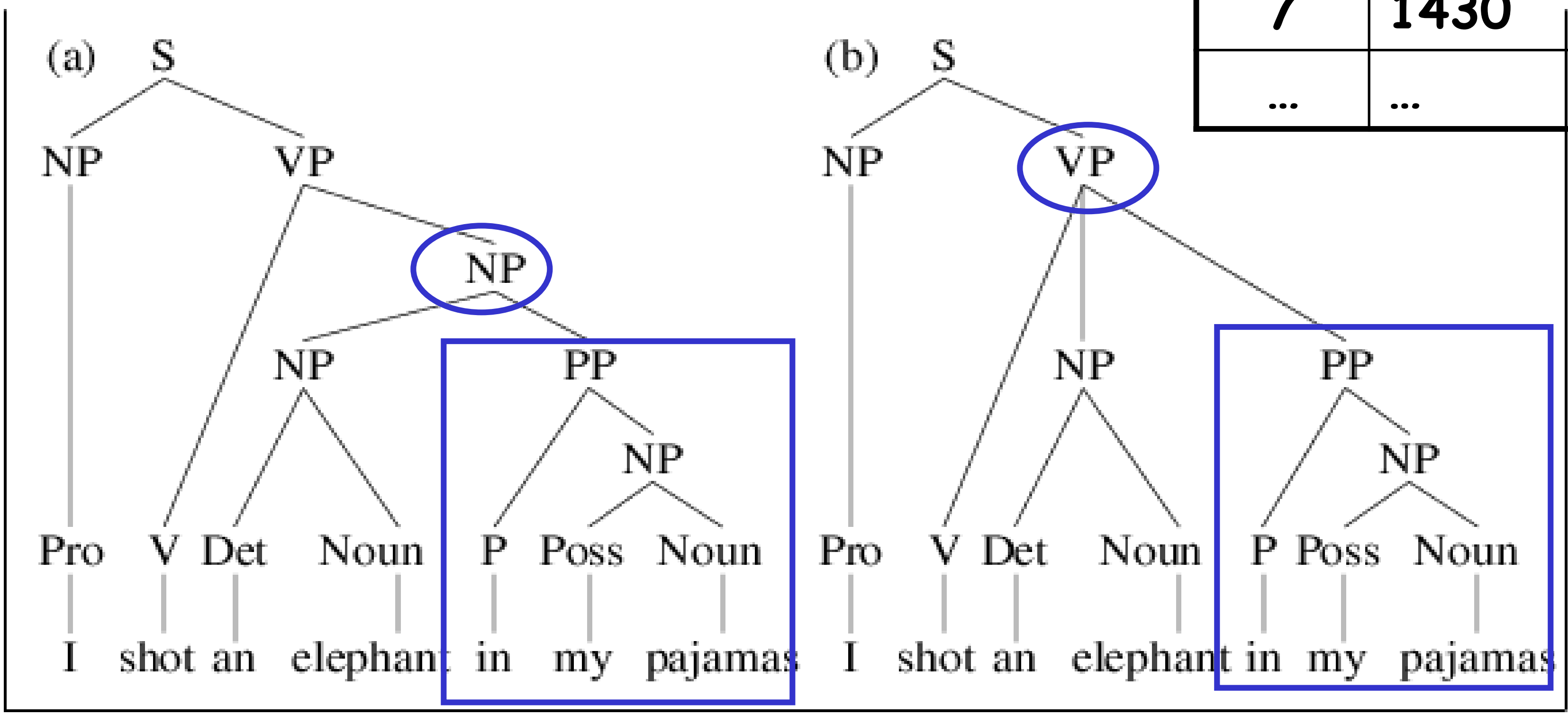
- Recap English Syntax and Parsing
- **Key Problem with parsing: Ambiguity**
- Probabilistic Context Free Grammars (PCFG)
- Treebanks and Grammar Learning

Structural Ambiguity (Ex. 1)

VP -> V NP ; NP -> NP PP
VP -> V NP PP

"I shot an elephant in my pajamas"

#of PP	# of NP parses
…	…
6	429
7	1430
…	…



Structural Ambiguity (Ex.2)

"I saw Mary passing by cs2"

"I saw Mary passing by cs2"

```
(ROOT
  (S
    (NP (PRP I))
    (VP (VBD saw)
      (S
        (NP (NNP Mary))
        (VP (VBG passing)
          (PP (IN by)
            (NP (NNP cs2)))))))
```

```
(ROOT
  (S
    (NP (PRP I))
    (VP (VBD saw)
        (NP (NNP Mary))
        (S
          (VP (VBG passing)
            (PP (IN by)
              (NP (NNP cs2)))))))
```

Structural Ambiguity (Ex. 3)

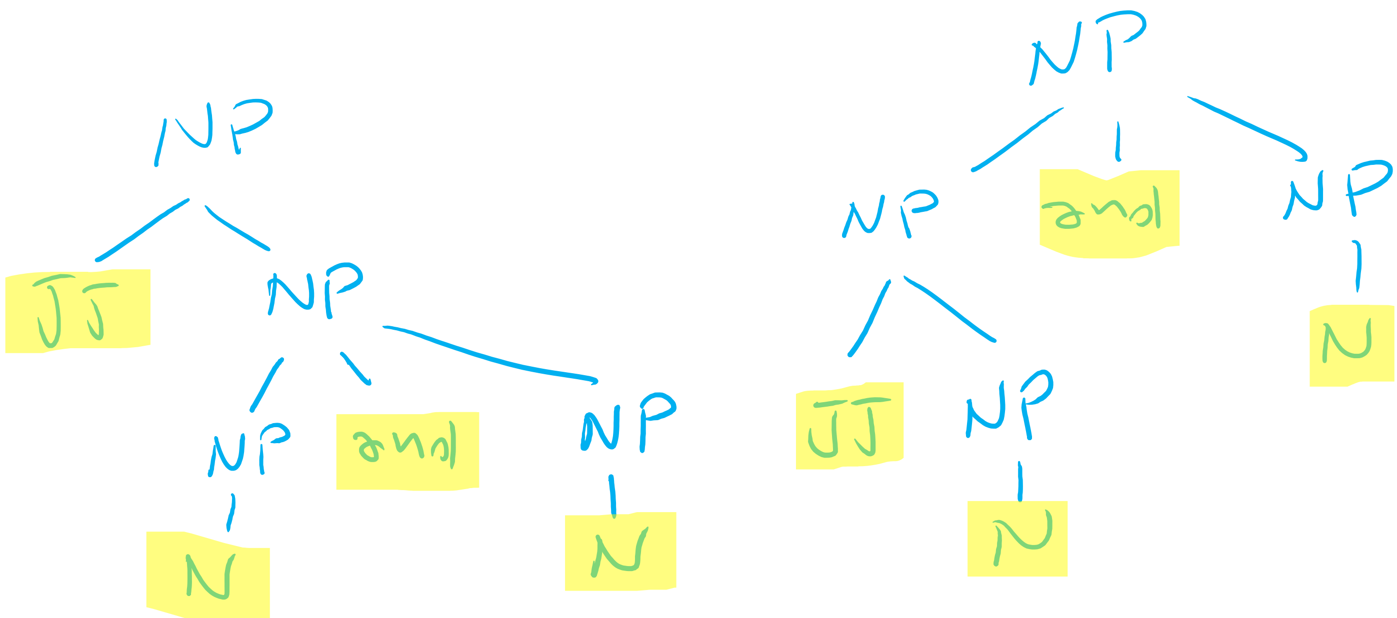
- **Coordination "new students and profs"**

JJ N N

NP -> NP and NP

NP -> JJ NP

NP -> N



Structural Ambiguity (Ex. 4)

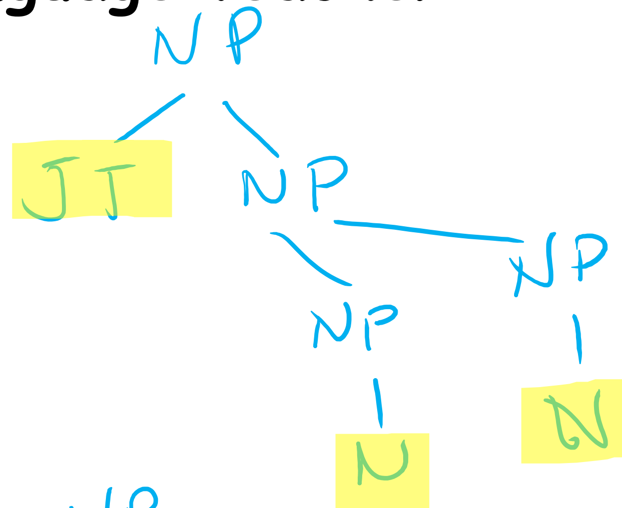
JJ N N

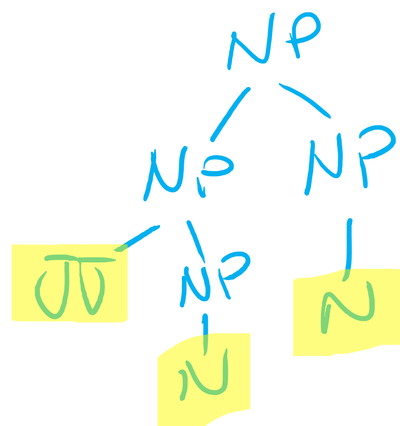
- **NP-bracketing "French language teacher"**

NP → JJ NP

NP → N

NP → NP NP





Approx. # of parses ?

- The famous Chinese cosmologists and astrophysicists saw a new spiral galaxy with the wide mirror telescope in Hawaii on the night of Jan 21st at 2:34 AM.

- Catalan of 5 = 14 *(5 prepositions)*
- Tot= 2 x 2 x 2 x 2 x 14 = 224

Lecture Overview

- Recap English Syntax and Parsing
- Key Problem with parsing: Ambiguity
- **Probabilistic Context Free Grammars (PCFG)**
- Treebanks and Grammar Learning (acquiring the probabilities)
- Intro to Parsing PCFG

Probabilistic CFGs (PCFGs)

- GOAL: assign a probability to parse trees and to sentences
- Each grammar rule is augmented with a conditional probability
- If these are all the rules for VP and .55 is P(VP->Verb | VP)

VP -> Verb	.55
VP -> Verb NP	.40
VP -> Verb NP NP	??

- What should ?? be ?

A. 1
B. 0
C. 0.05
D. 0.42
E. None of the above

Probabilistic CFGs (PCFGs)

- **GOAL: assign a probability to parse trees and to sentences**
- **Each grammar rule is augmented with a conditional probability**
  - **The expansions for a given non-terminal sum to 1**

VP -> Verb	.55	P(VP->Verb | VP)
VP -> Verb NP	.40	P(VP->Verb NP | VP)
VP -> Verb NP NP	.05	P(VP->Verb NP NP | VP)

Formal Def: 5-tuple (N, ∑, P, S, D)

Sample PCFG

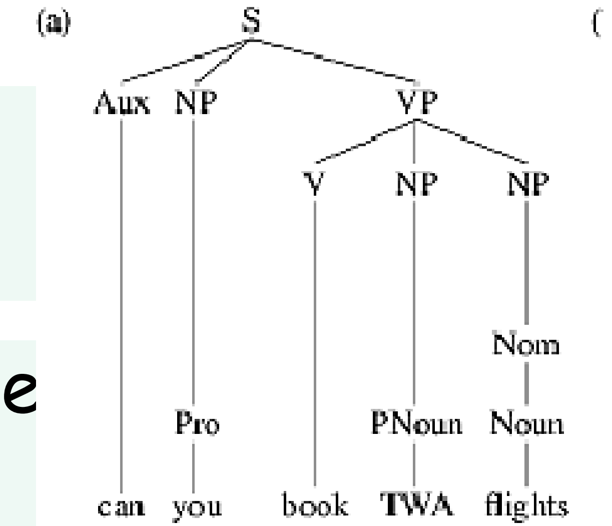
S → *NP VP*	[.80]	*Det* → *that* [.05] | *the* [.80] | *a* [.15]	
S → *Aux NP VP*	[.15]	*Noun* → *book*	[.10]
S → *VP*	[.05]	*Noun* → *flights*	[.50]
NP → *Det Nom*	[.20]	*Noun* → *meal*	[.40]
NP → *Proper-Noun*	[.35]	*Verb* → *book*	[.30]
NP → *Nom*	[.05]	*Verb* → *include*	[.30]
NP → *Pronoun*	[.40]	*Verb* → *want*	[.40]
Nom → *Noun*	[.75]	*Aux* → *can*	[.40]
Nom → *Noun Nom*	[.20]	*Aux* → *does*	[.30]
Nom → *Proper-Noun Nom*	[.05]	*Aux* → *do*	[.30]
VP → *Verb*	[.55]	*Proper-Noun* → *TWA*	[.40]
VP → *Verb NP*	[.40]	*Proper-Noun* → *Denver*	[.40]
VP → *Verb NP NP*	[.05]	*Pronoun* → *you* [.40] | *I* [.60]	

PCFGs are used to….

- **Estimate Prob. of parse tree**

A. Sum of the probs of all the rules applied

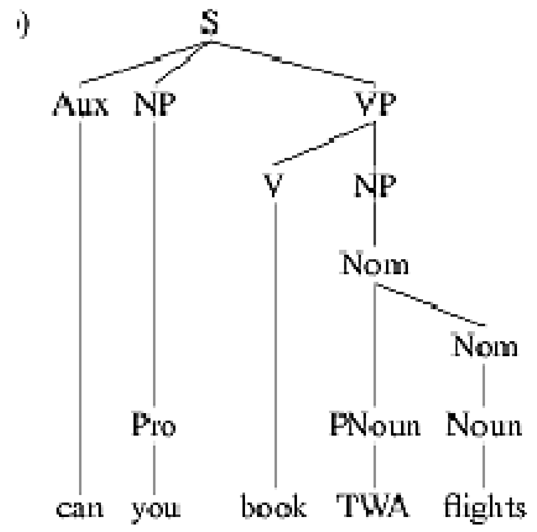
B. Product of the probs of all the rules applied



- **Estimate Prob. of a sentence**

A. Sum of the probs of all the parse trees

B. Product of the probs of all the parse trees



PCFGs are used to….

- **Estimate Prob. of parse tree**

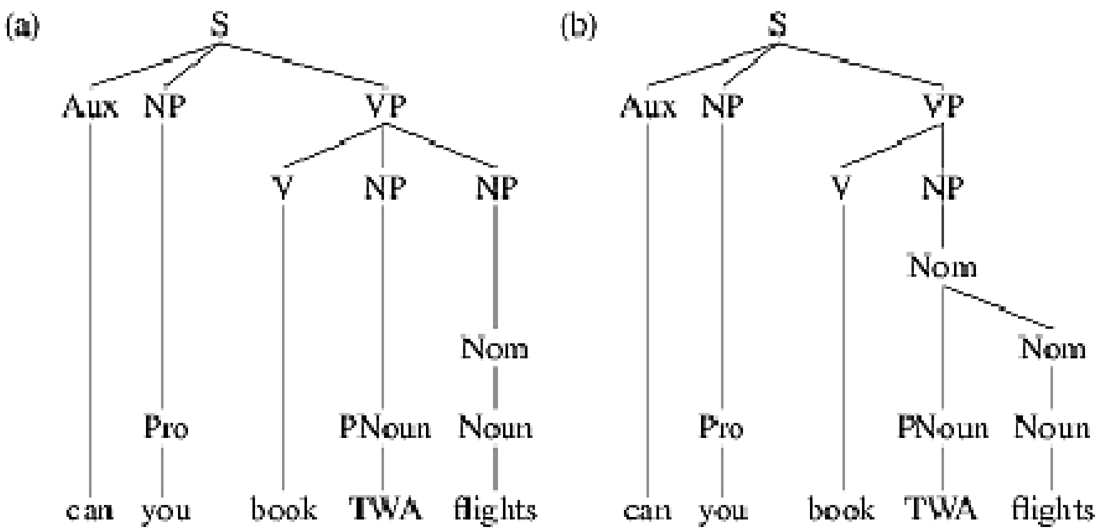
$$P(Tree) = \prod_{node \in Tree} P(\text{expansion for node})$$

- **Estimate Prob. to sentences**

$$P(Sentence) = \sum_{Tree \in \text{Sentence parses}} P(Tree)$$

Example

$P(Tree^a) = .15 * .4 * \ldots = 3.8 \cdot 10^{-7}$ $\quad P(Tree^b) = .15 * .4 * \ldots = 4.3 \cdot 10^{-7}$



P("can you book TWA flights")

$8.1 \cdot 10^{-7}$

(a) →

Rules			P	Rules			P
S	→	Aux NP VP	.15	S	→	Aux NP VP	.15
NP	→	Pro	.40	NP	→	Pro	.40
VP	→	V NP NP	.05	VP	→	V NP	.40
NP	→	Nom	.05	NP	→	Nom	.05
NP	→	PNoun	.35	Nom	→	PNoun Nom	.05
Nom	→	Noun	.75	Nom	→	Noun	.75
Aux	→	Can	.40	Aux	→	Can	.40
~~NP~~	~~→~~	~~Pro~~	.40	~~NP~~	~~→~~	~~Pro~~	.40
Pro	→	you	.40	Pro	→	you	.40
Verb	→	book	.30	Verb	→	book	.30
PNoun	→	TWA	.40	Pnoun	→	TWA	.40
Noun	→	flights	.50	Noun	→	flights	.50

← (b)

Lecture Overview

- Recap English Syntax and Parsing
- Key Problem with parsing: Ambiguity
- Probabilistic Context Free Grammars (PCFG)
- **Treebanks and Grammar Learning (acquiring the probabilities)**

Treebanks

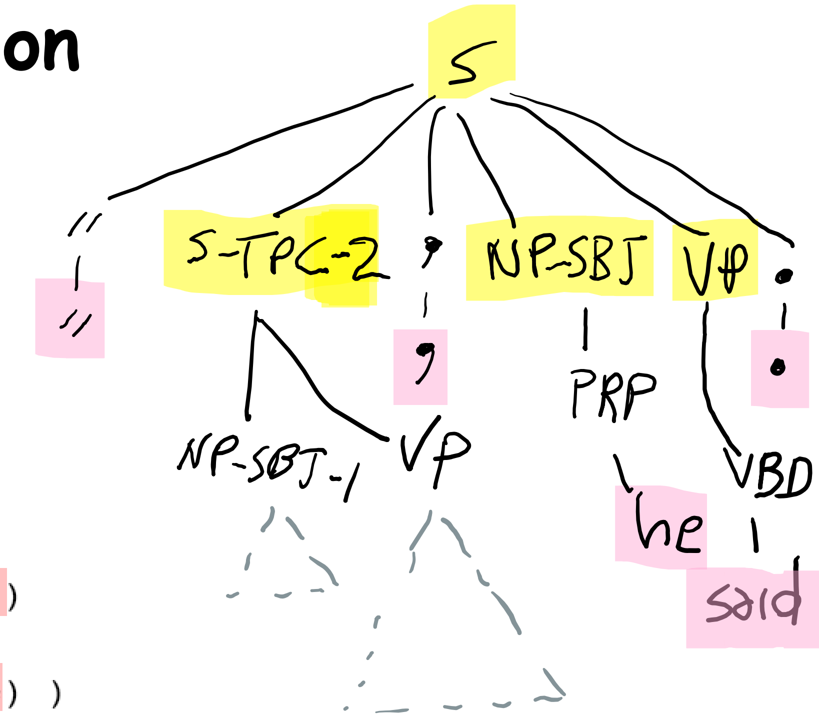
- Definition: corpora in which each sentence has been paired with a parse tree
- These are generally created
  - Parse collection with parser
  - human annotators revise each parse
- Requires detailed annotation guidelines
  - POS tagset
  - Grammar
  - instructions for how to deal with particular grammatical constructions.

Treebanks

- **Definition: corpora in which each sentence has been paired with a parse tree**
- **These are generally created**
  - **By human annotators**
- **Requires detailed annotation guidelines**
  - **POS tagset**
  - **Grammar**
  - **instructions for how to deal with particular grammatical constructions.**

Penn Treebank

- **Penn TreeBank is a widely used treebank. Trees are in bracket notation**



- Most well known is the Wall Street Journal section of the Penn TreeBank.
  - 1 M words from the 1987-1989 Wall Street Journal.

```
( (S (`` ``)
    (S-TPC-2
      (NP-SBJ-1 (PRP We) )
      (VP (MD would)
        (VP (VB have)
          (S
            (NP-SBJ (-NONE- *-1) )
            (VP (TO to)
              (VP (VB wait)
                (SBAR-TMP (IN until)
                  (S
                    (NP-SBJ (PRP we) )
                    (VP (VBP have)
                      (VP (VBN collected)
                        (PP-CLR (IN on)
                          (NP (DT those)(NNS assets)))))))))))))
    (, ,) ('' '')
    (NP-SBJ (PRP he) )
    (VP (VBD said)
      (S (-NONE- *T*-2) ))
    (. .) ))
```

1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential *there*
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRP$	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb, comparative
22.	RBS	Adverb, superlative
23.	RP	Particle
24.	SYM	Symbol
25.	TO	*to*
26.	UH	Interjection
27.	VB	Verb, base form
28.	VBD	Verb, past tense
29.	VBG	Verb, gerund or present participle
30.	VBN	Verb, past participle
31.	VBP	Verb, non-3rd person singular present
32.	VBZ	Verb, 3rd person singular present
33.	WDT	Wh-determiner
34.	WP	Wh-pronoun
35.	WP$	Possessive wh-pronoun
36.	WRB	Wh-adverb

Treebank Grammars

- Such grammars tend to contain lots of rules….
- For example, the Penn Treebank has 4500 different rules for VPs! Among them...

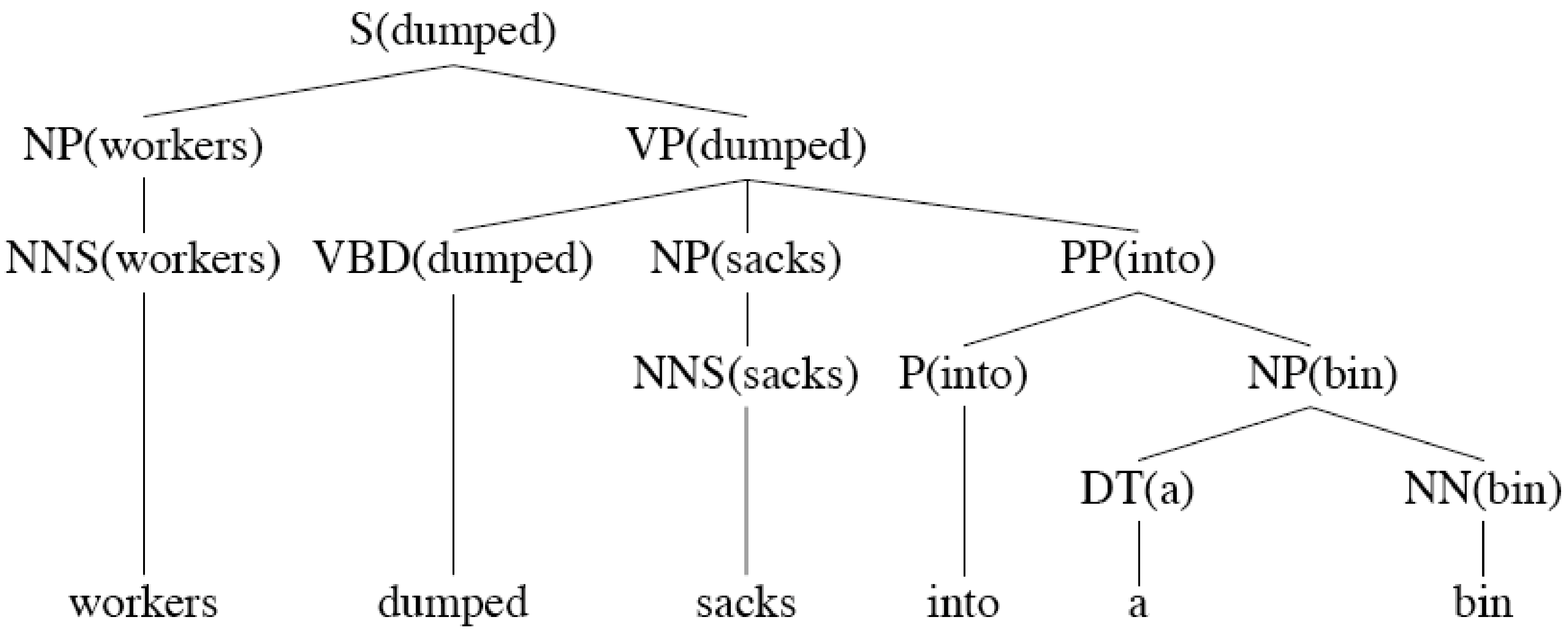
$$VP \rightarrow VBD\ PP$$
$$VP \rightarrow VBD\ PP\ PP$$
$$VP \rightarrow VBD\ PP\ PP\ PP$$
$$VP \rightarrow VBD\ PP\ PP\ PP\ PP$$

Heads in Trees

- Finding heads in treebank trees is a task that arises frequently in many applications.
  - Particularly important in statistical parsing (e.g., with PCFG)

- We can visualize this task by annotating the nodes of a parse tree with the heads of each corresponding node.

Lexically Decorated Tree

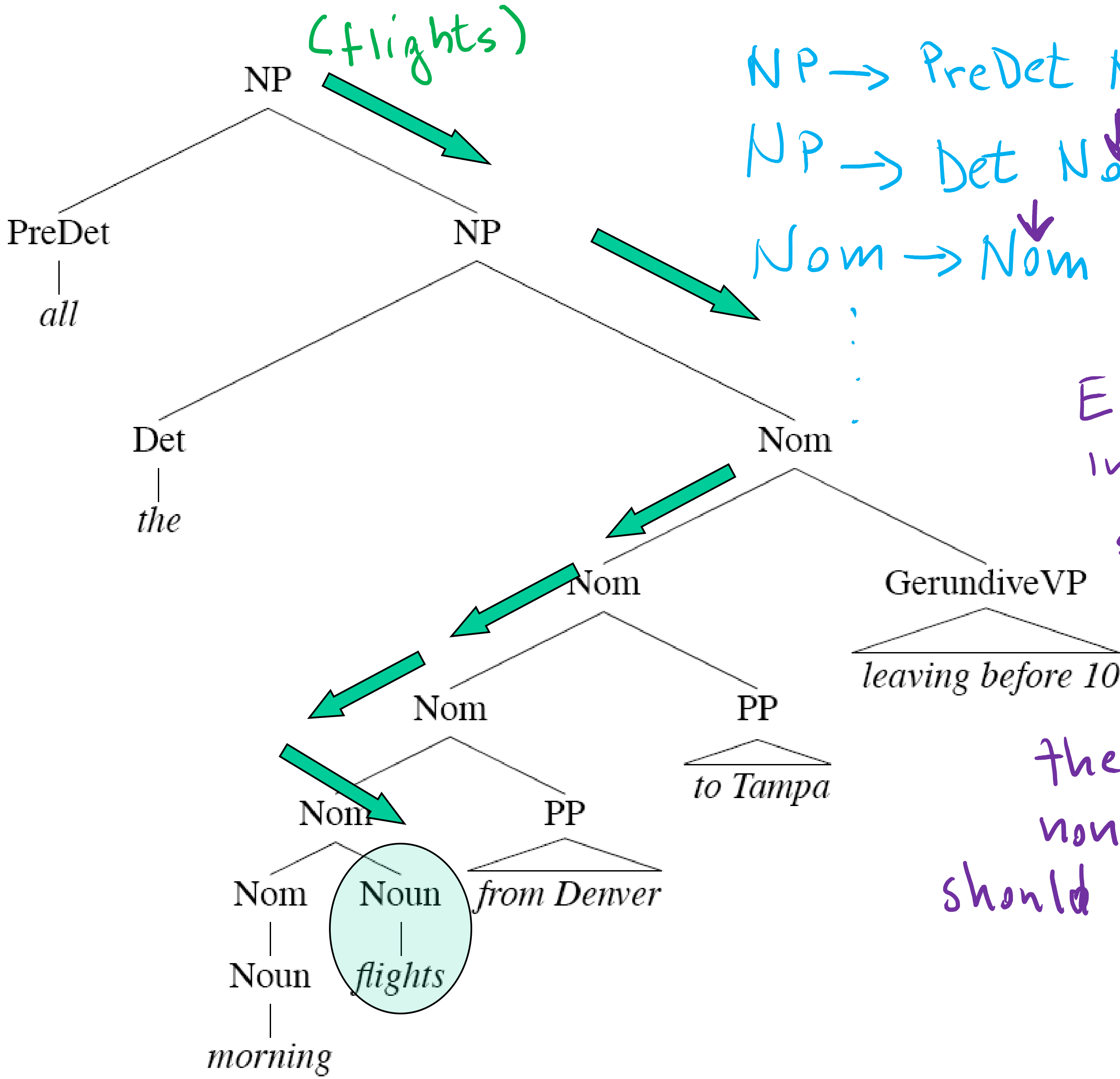


Head Finding

- **The standard way to do head finding is to use a simple set of tree traversal rules specific to each non-terminal in the grammar.**

- **Each rule in the PCFG specifies where the head of the expanded non-terminal should be found**

Noun Phrases



NP → PreDet NP

NP → Det Nom

Nom → Nom GerundiveVP

Each rule in the PCFG specifies where the head of the expanded non-terminal should be found

Acquiring Grammars and Probabilities

Manually parsed text corpora (e.g., PennTreebank)

- **Grammar**: read it off the parse trees

Ex: if an NP contains an ART, ADJ, and NOUN then we create the rule *NP -> ART ADJ NOUN*.

- **Probabilities:**

$$P(A \to \alpha \mid A) = \frac{count(A \to \alpha)}{\sum_{\forall \gamma} count(A \to \gamma)} = \frac{count(A \to \alpha)}{count(A)}$$

Ex: if the *NP -> ART ADJ NOUN* rule is used 50 times and all NP rules are used 5000 times, then the rule's probability is … .01

Example

If you look at all the parse trees in the bank you find three rules for NP

Rule	How many times
(1) NP → ART ADJ NOUN	50
(2) NP → NOUN	4000
(3) NP → PRONOUN	950
	5000 total # of NP expansions

$P(① | NP) = 50/5000 = .01$

$P(② | NP) = 4000/5000 = .8$

$P(③ | NP) = 950/5000 = .19$

$.19 = 1 - (.01 + .8)$

Learning Goals for today's class

You can:

- Provide a formal definition of a PCFG
- Apply a PCFG to compute the probability of a parse tree of a sentence as well as the probability of a sentence
- Describe the content of a treebank
- Describe the process to identify a head of a syntactic constituent
- Compute the probability distribution of a PCFG from a treebank

Next class Mon (March 22)

- Parsing Probabilistic CFG: CKY parsing
- PCFG in practice: Modeling Structural and Lexical Dependencies

Keep working on Assignment-3 - due Mar 30

(8-18 hours - working in pairs on programming parts is strongly advised)